Visual Odometry

Yang Cheng
Machine Vision Group
Section 348
ycheng@jpl.nasa.gov Phone: 4-1857
12/16/2003

Outline of this Talk

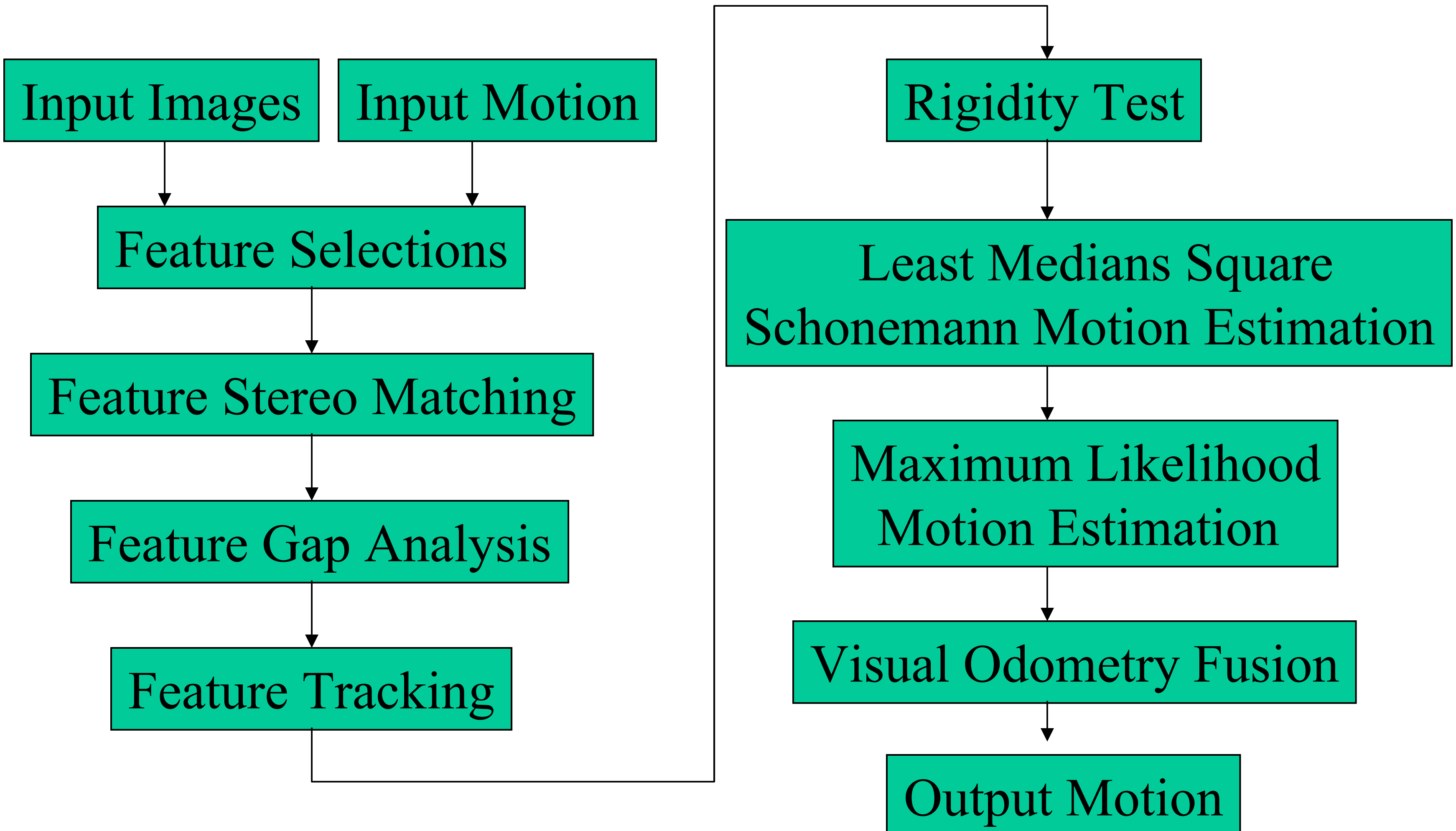
- Brief History
- Algorithm
- Software structure and interface
- Software Features
- Ground truth measurement
- Some results
- Future works

Brief History

- H. Moravec's PhD Thesis, " Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover, Stanford University, 1980
- Larry Matthies' PhD Thesis, "Dynamic Stereo Vision", Oct, 1989, CMU.
- A version of Visual Odometry in C was implemented in early 1990s in JPL.
- A C++ version of visual odometry was implemented by MTP Slope Navigation task led by Larry Matthies in 2001.
- The visual odometry has been ported to CLARAty and demonstrated onboard motion estimation on Rock 8 in 2002.
- The visual odometry has been used successfully on slip compensation by the slope navigation task.
- The visual odometry has been integrated officially to MER navigation software and demonstrated successfully in 2003.
- A few other versions of visual odometry were developed in academic and industry communities.

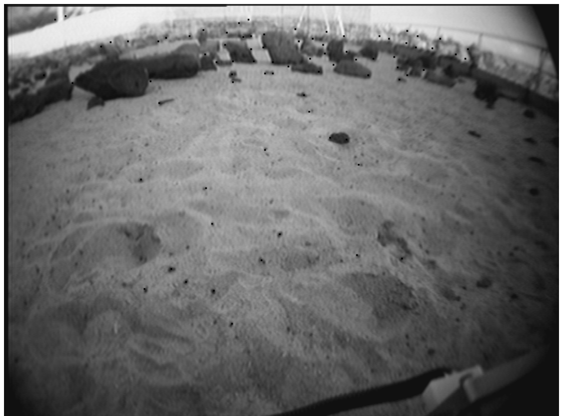
Visual Odometry

To use a (stereo) image sequence to track 3-D point features, or landmark, to estimate the motion of the vehicle.

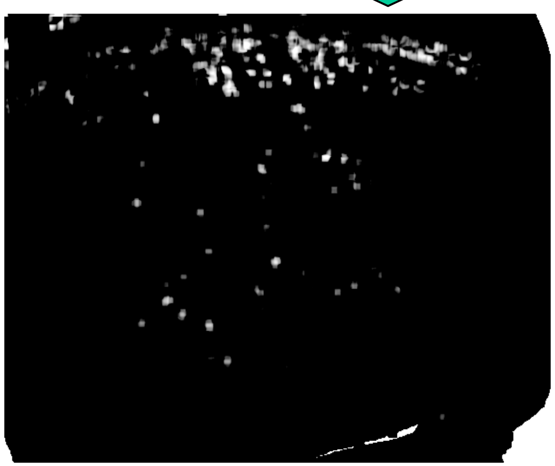


Feature (Landmark) Selection

Input Image

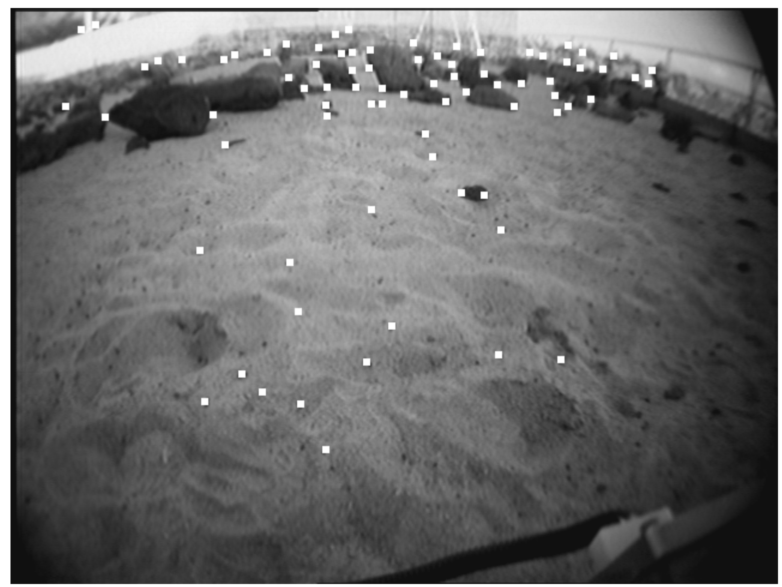


Forstner operator



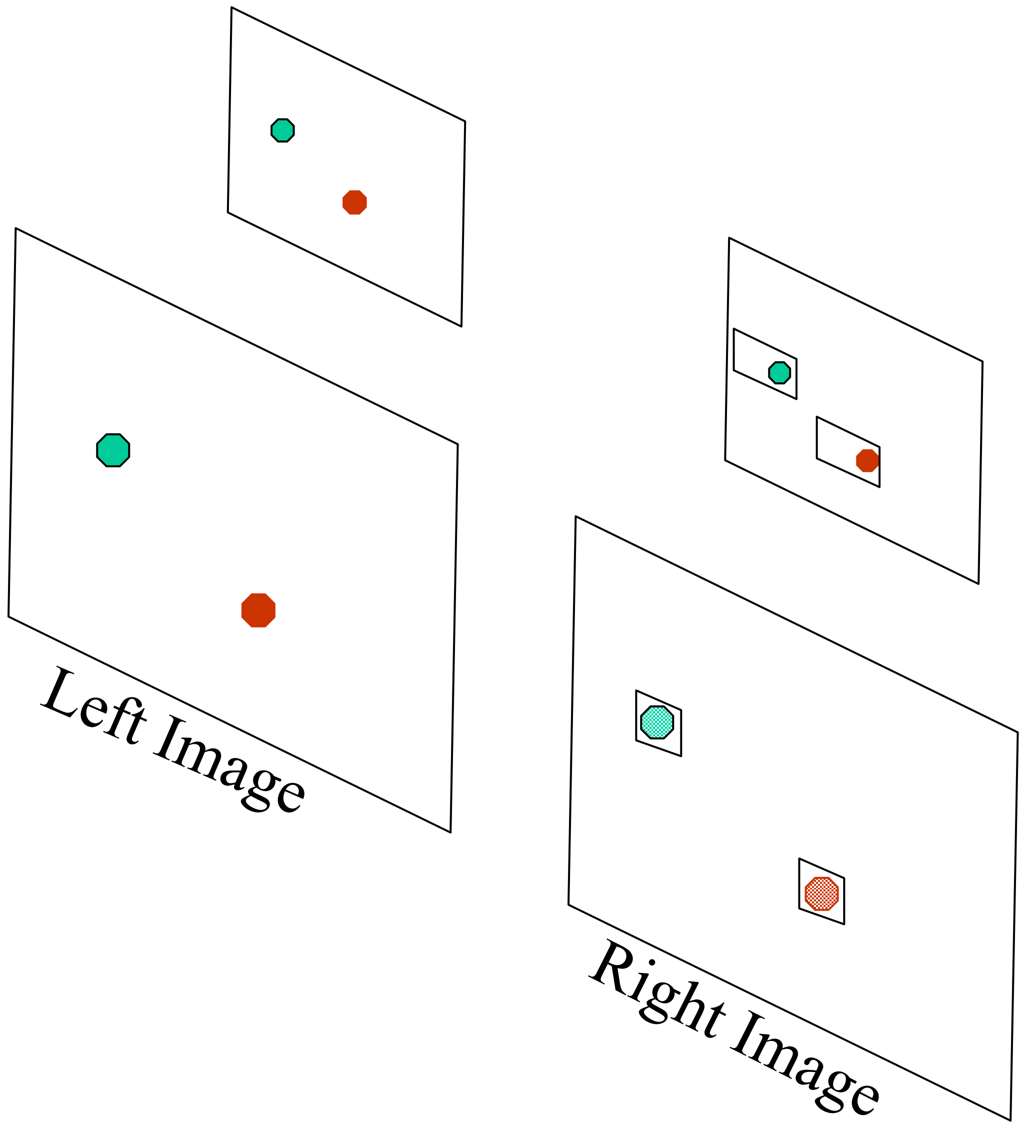
Interest Image

- **A landmark is a patch of image which must exhibit intensity variation that allows the landmark to be localized in subsequent image.**

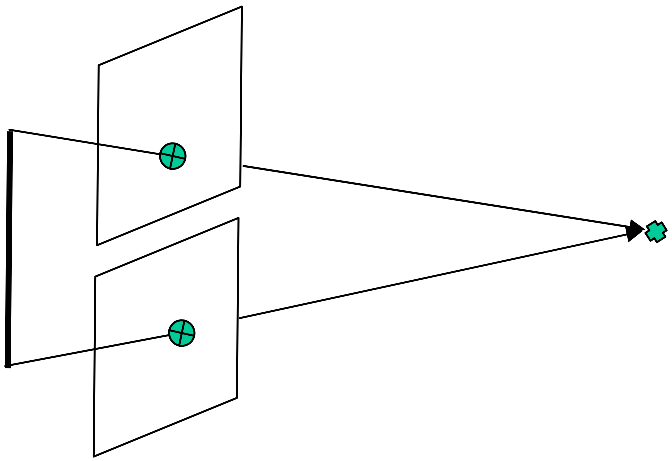


Landmarks

Feature Stereo Matching (Pyramid Searching)

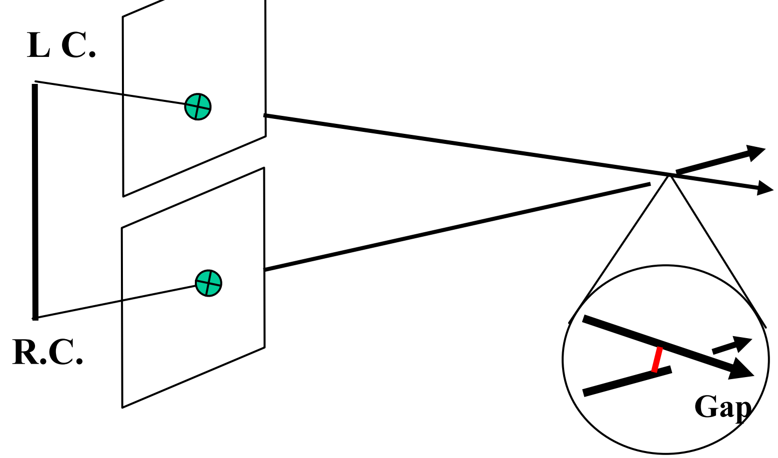


Feature Gap Analysis and Triangulation Error

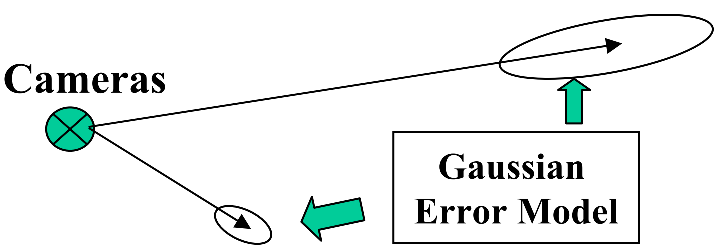


Gap Analysis

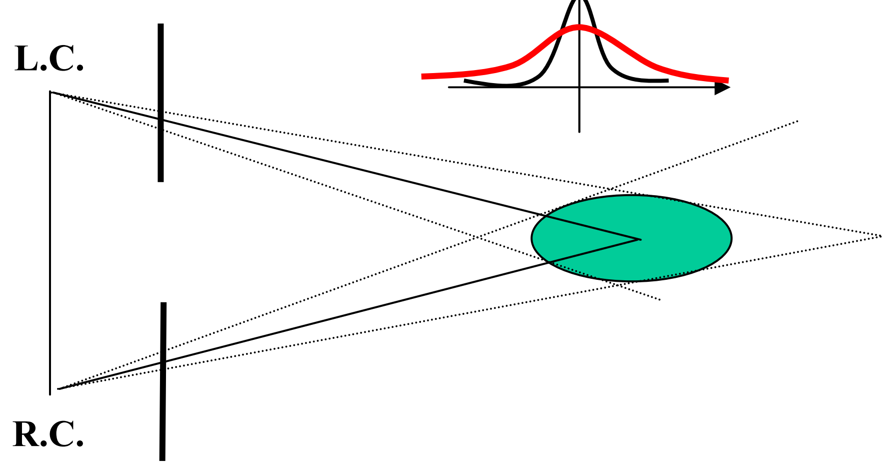
Error Vs Ray Gap



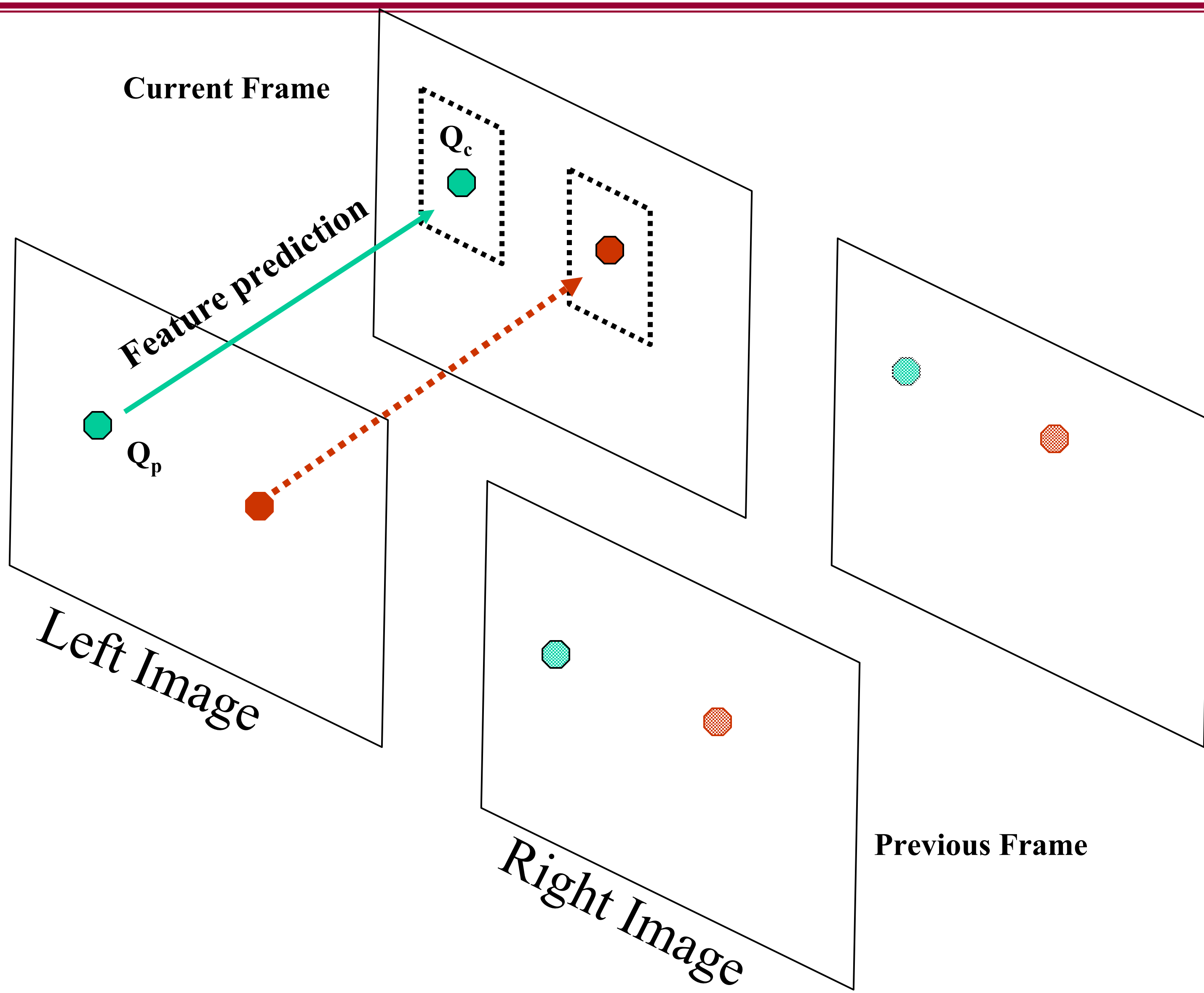
Error Vs Location



Error Vs Correlation Peak



Feature Tracking



Motion Estimation (Least-Squares Vs Maximum Likelihood)

$$Q_{ci} = RQ_{pi} + T + v_i$$

$$e_i = Q_c - RQ_{pi} - T$$

$$q(R,T) = \sum_i w_i e_i^T e_i$$

Least-squares Estimation

- A closed form solution
- Rotation, *R,* with orthogonal constrain is estimated first
- Translation, *T*, is then estimated.
- Reflect the quality of the observations.
- It is fast.
- The resulting motion estimates can be substantially inferior.

Maximum Likelihood Estimation

- An nonlinear optimization solution
- Fully reflects the error model
- It is relative slow
- It needs an initial estimate.
- It is sensitive to outliers
- Its motion estimates in general is much superior than the least-squares estimation.

Least-squares Estimation

Merit Function:

$$e_i = Q_c - RQ_{pi} - T$$

$$q(R,T) = \sum_i w_i e_i^T e_i$$

Orthogonal constrains:

$$r_i^T r_i = 1 \qquad r_i^T r_j = 0 \qquad i,j \in \{1,2,3\} i \neq j$$

$$q(R,T,l_i,m_i) = \{\sum w_i e_j^T e_j\} + \sum_{i=1}^{3} l_i(r_i^T r_i - 1) + \sum_{i,j=0}^{3} m_i r_i^T r_j$$

Solutions:

$$w = \sum w_i \quad Q_1 = \sum Q_{ci} \quad Q_2 = \sum Q_{pi}$$

$$A = \sum w_i Q_{pi} Q_{ci}^T \quad E = A - \frac{1}{w} Q_1 Q_2^T \quad E = USV^T$$

$$R = UV^T \quad T = \frac{1}{w}[Q_1 - RQ_2]$$

Maximum Likelihood Estimation

Merit Function:

$$M = \sum e_i^T W e_i$$

W = covariance matrix of the feature i

Solutions: To linearize the merit function and determine the three attitude and three translation iteratively.
Page 150 of Larry Matthies' thesis

Visual Odometry Interface

VOMotionStart(leftCam, rightCam, ParameterFile, leftImage, rightImage, leftDisp, InitialMotion)

VOMotion(leftImage, rightImage, leftDisp, InitialMotion, *estMotion)

Camera models: CAHV, CAHVOR, CAHVORE

leftDisp: the disparity image generated by stereo processing.

Motion file: Position[3], attitude [3], covarence[6][6]

Parameter File contains 48 parameters

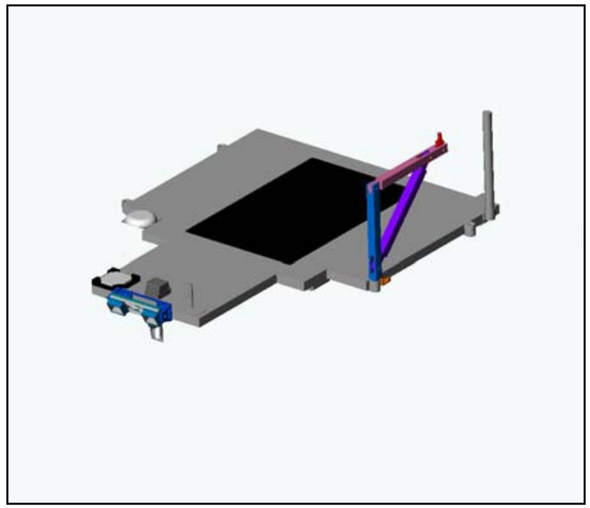
Some VO Parameters

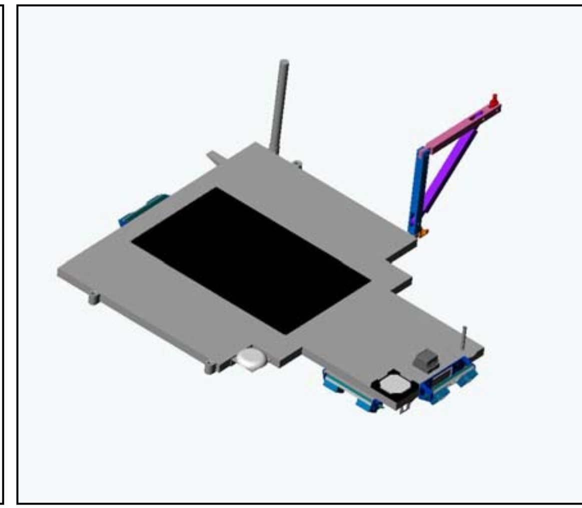
VO_MAX_NUM_VO_FEATURES	600	features
VO_MIN_NUM_VO_FEATURES	8	iteration
VO_VO_MAX_PIXEL_OFFSET	1	pixel
VO_MAX_VO_ITERATIONS	50	iteration
VO_VO_CORR_WINDOW_ROWS	9	pixel
VO_VO_CORR_WINDOW_COLS	9	pixel
VO_VO_TRACK_WINDOW_SIZE	50	pixel
VO_VO_SELECT_WINDOW_SIZE	9	pixel
VO_VO_NUM_IMAGE_PAIRS	4	images
VO_VO_IMAGE_ROWS	640	pixel
VO_VO_IMAGE_COLS	480	pixel
VO_SCHONEMANN_ITERATIONS	50	iteration
VO_VO_MIN_DIST_FEATURE	0.5	meter
VO_VO_MAX_DIST_FEATURE	20.0	meter
VO_VO_AFFINE_MATCH_FLAG	0	Boolean
VO_MAX_DELTA	0.000006	
VO_DEFAULT_VO_MIN_CORRELATION	0.8	correlation

Arroyo Data Collection

1. About 8 meters of image (20 cm step) sequence were collected at JPL arroyo in March, 2002.
2. Onboard IMU, wheel odometry and other data were collected.
3. Ground truth data (position and attitude) were collected by totalstation.

Total Station & Prism

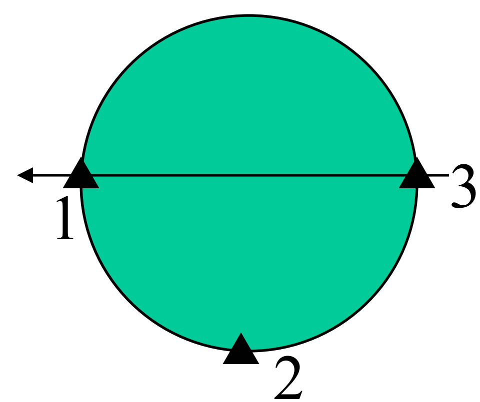




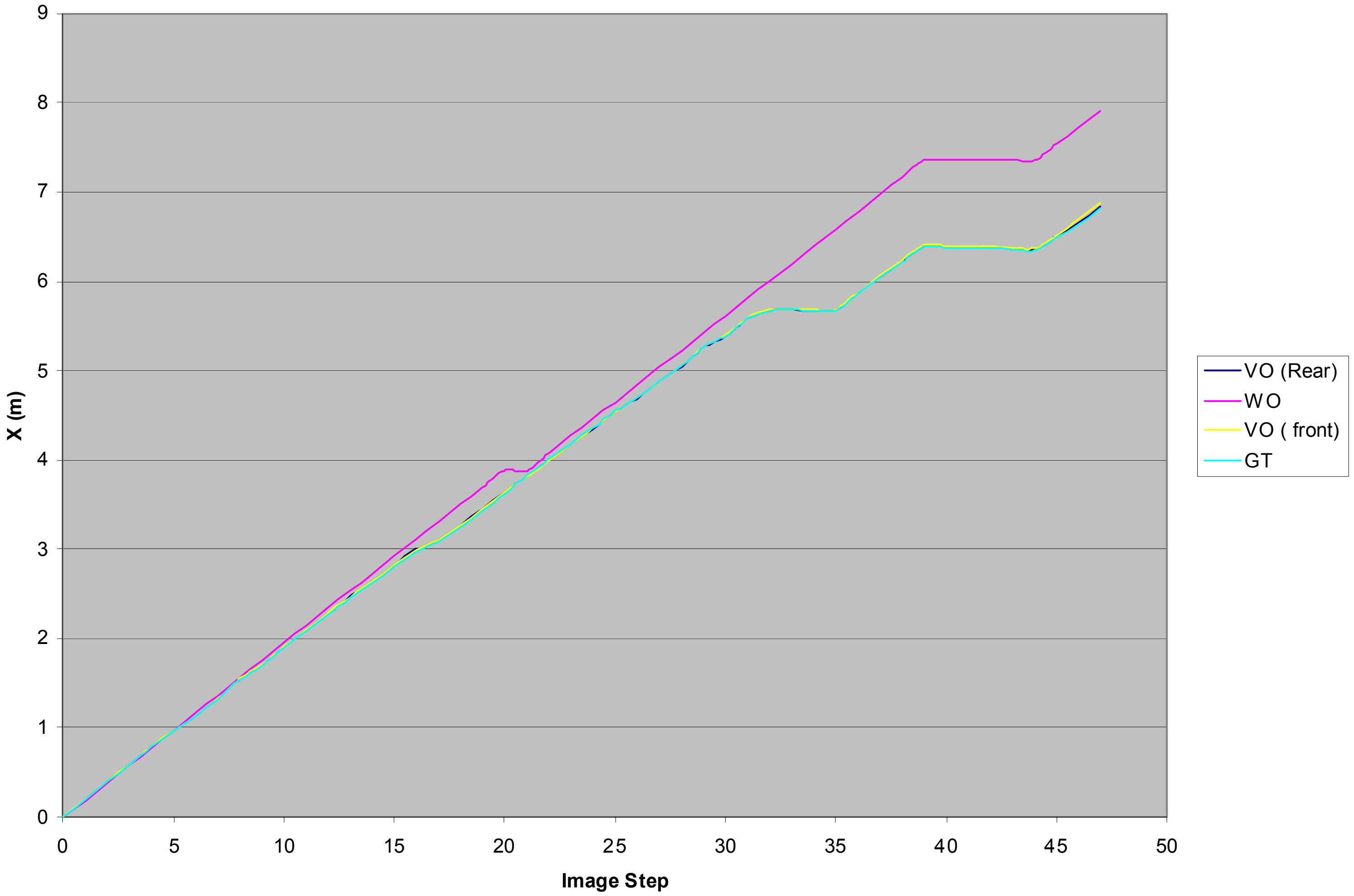
Three points are measured at each stop.

The position and attitude can be determined.

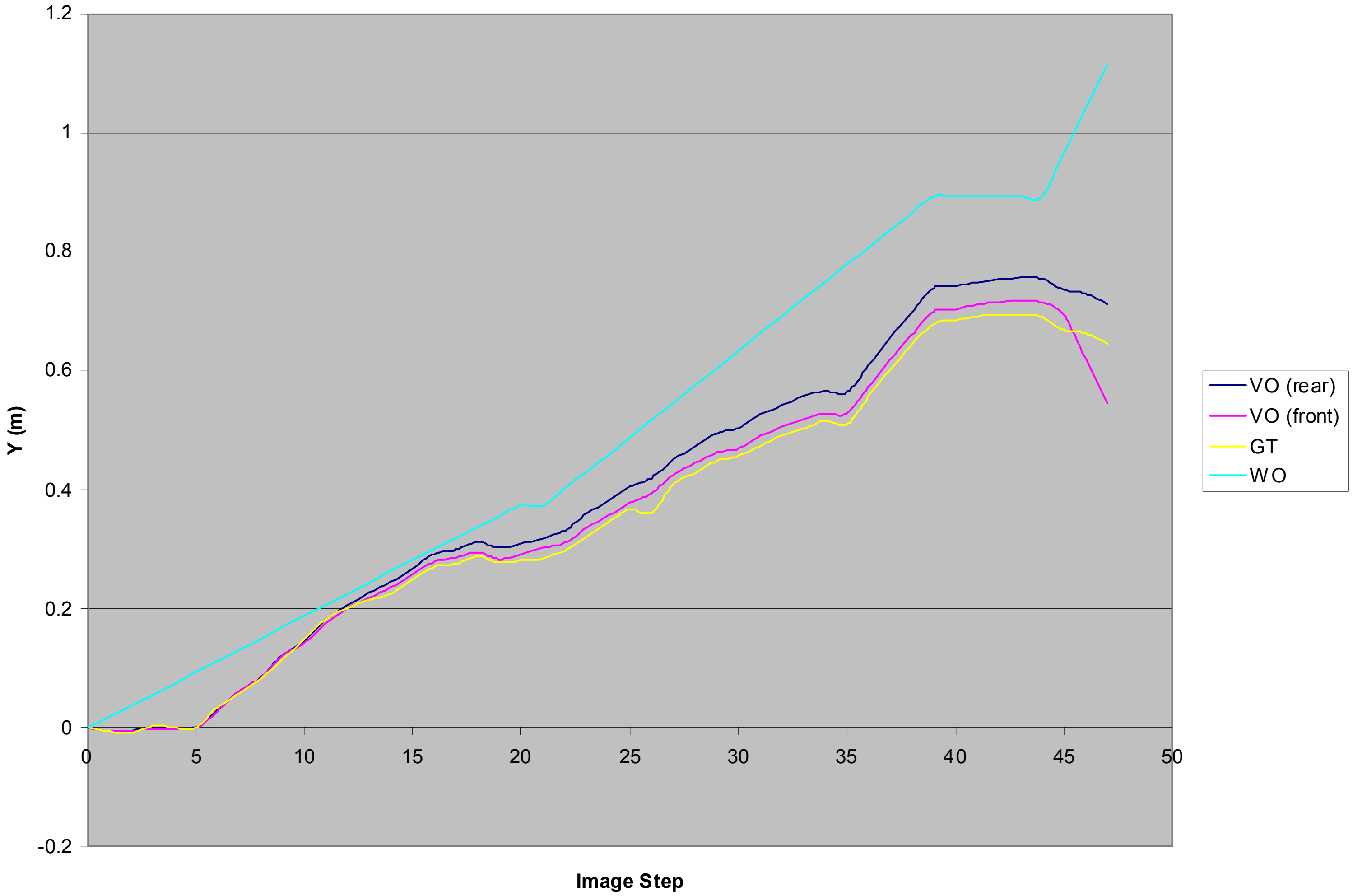
Pitch, Roll, Heading error < 0.5 degree;
Position error < 3mm.



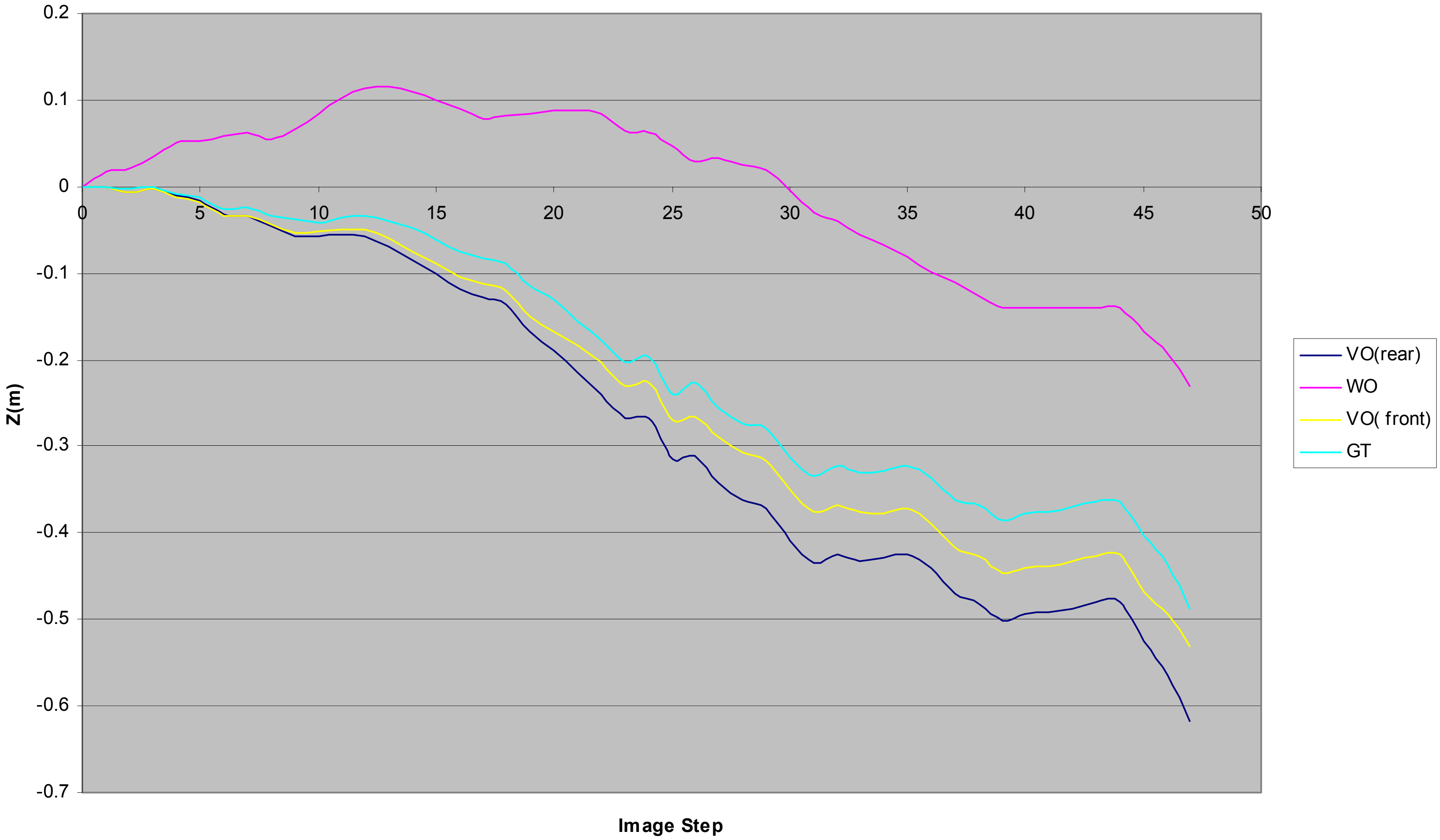
Motion Estimation (X)



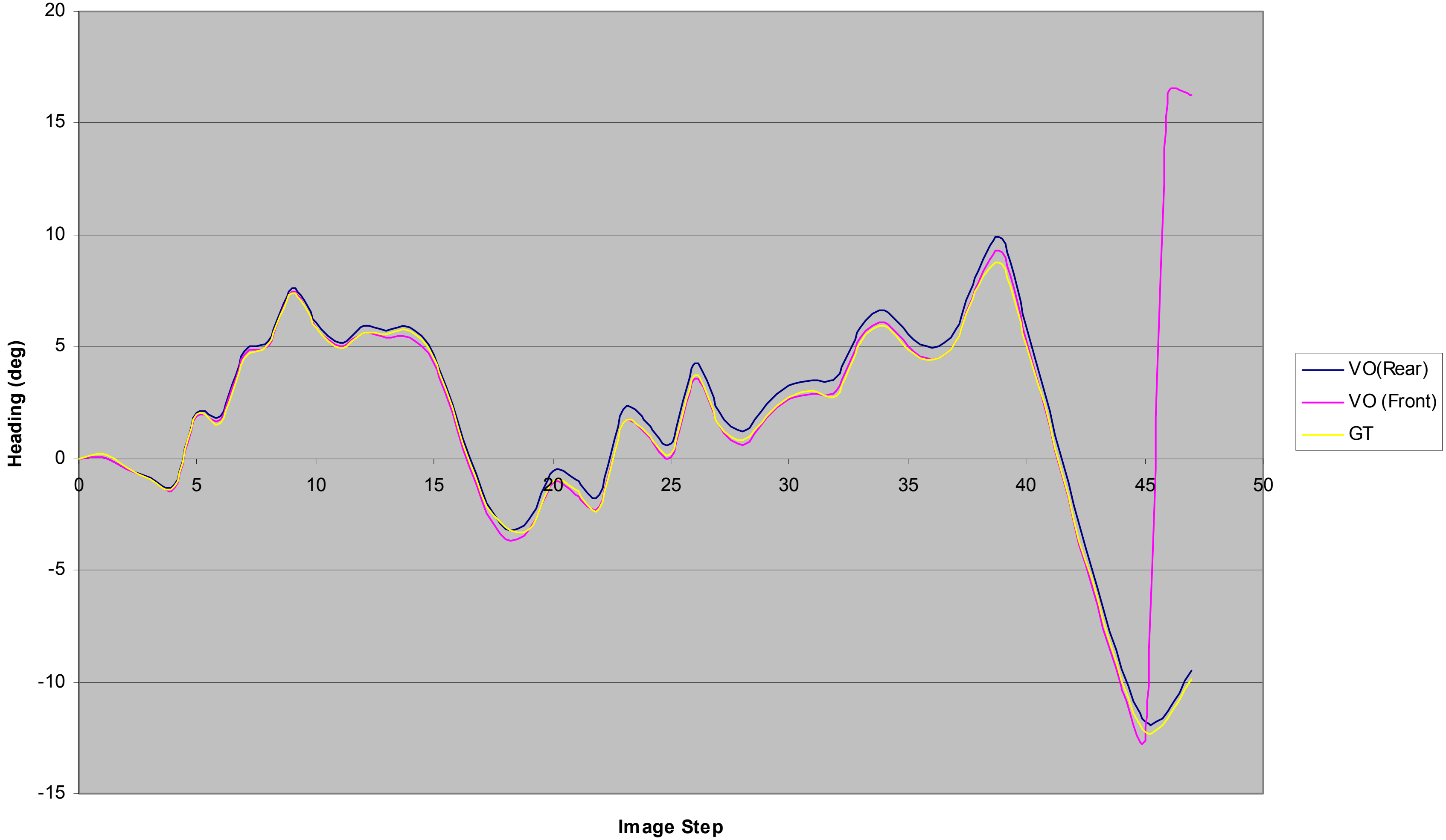
Motion Estimation (Y)



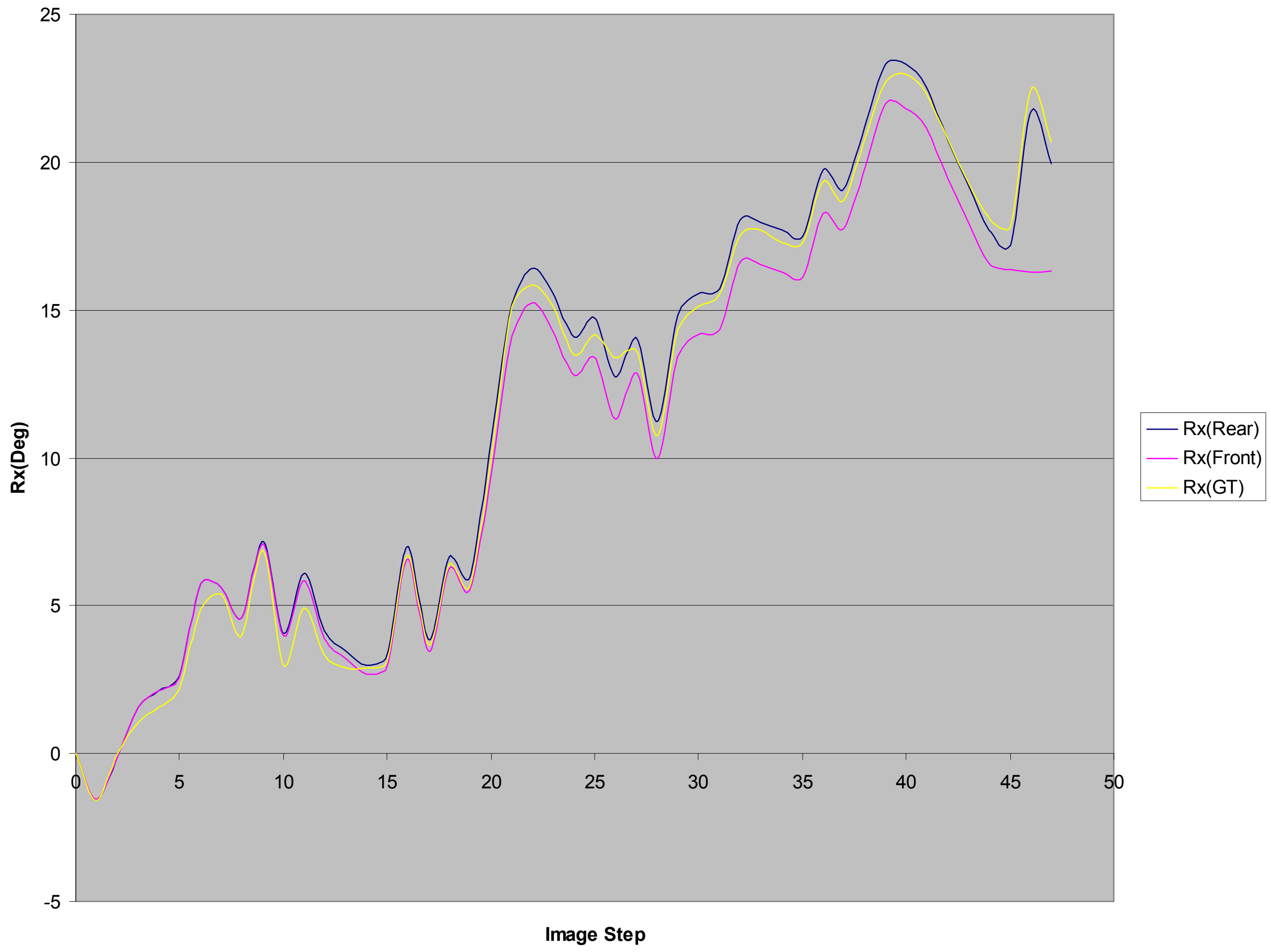
Motion Estimation (Z)



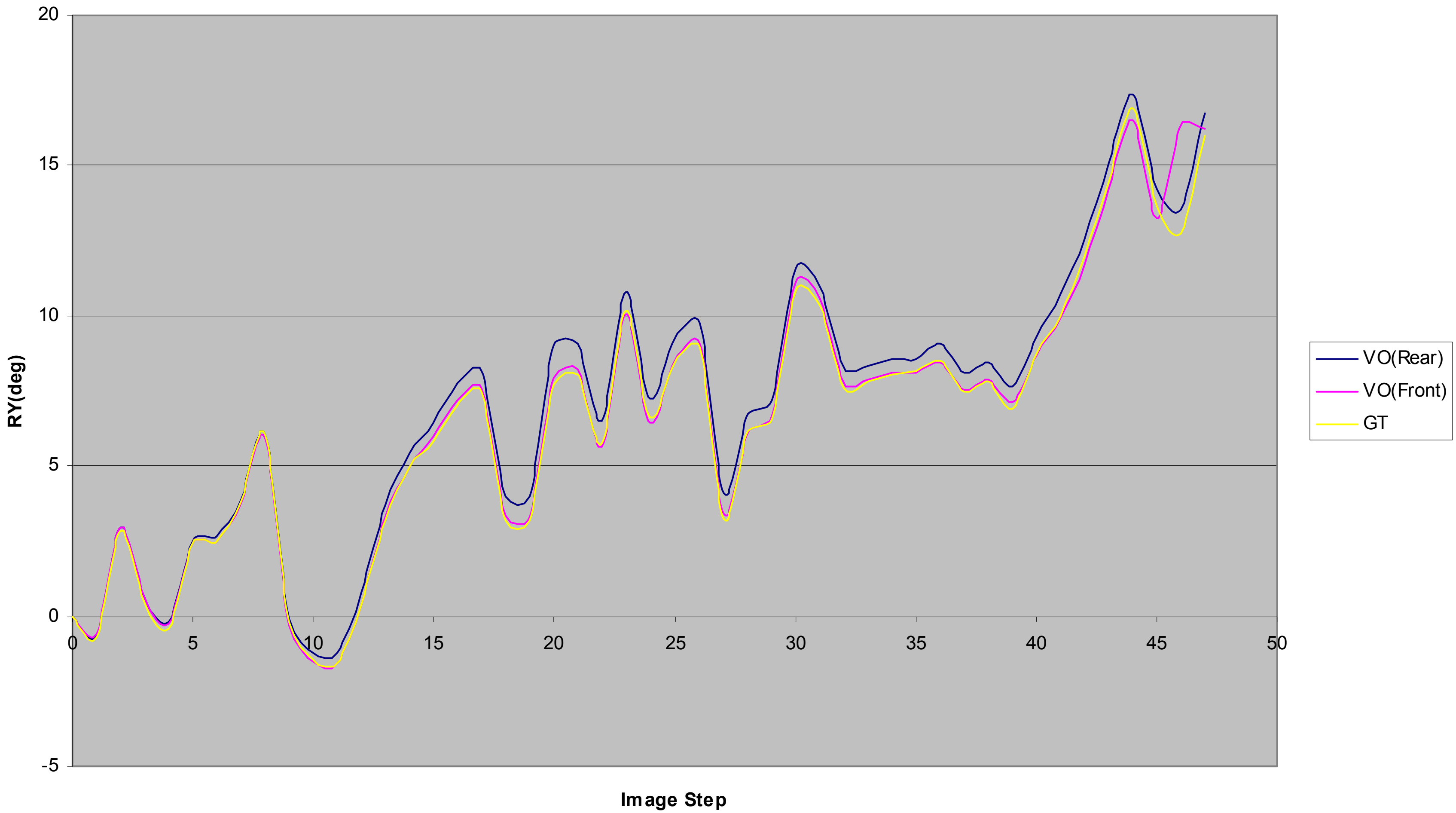
Heading Estimation



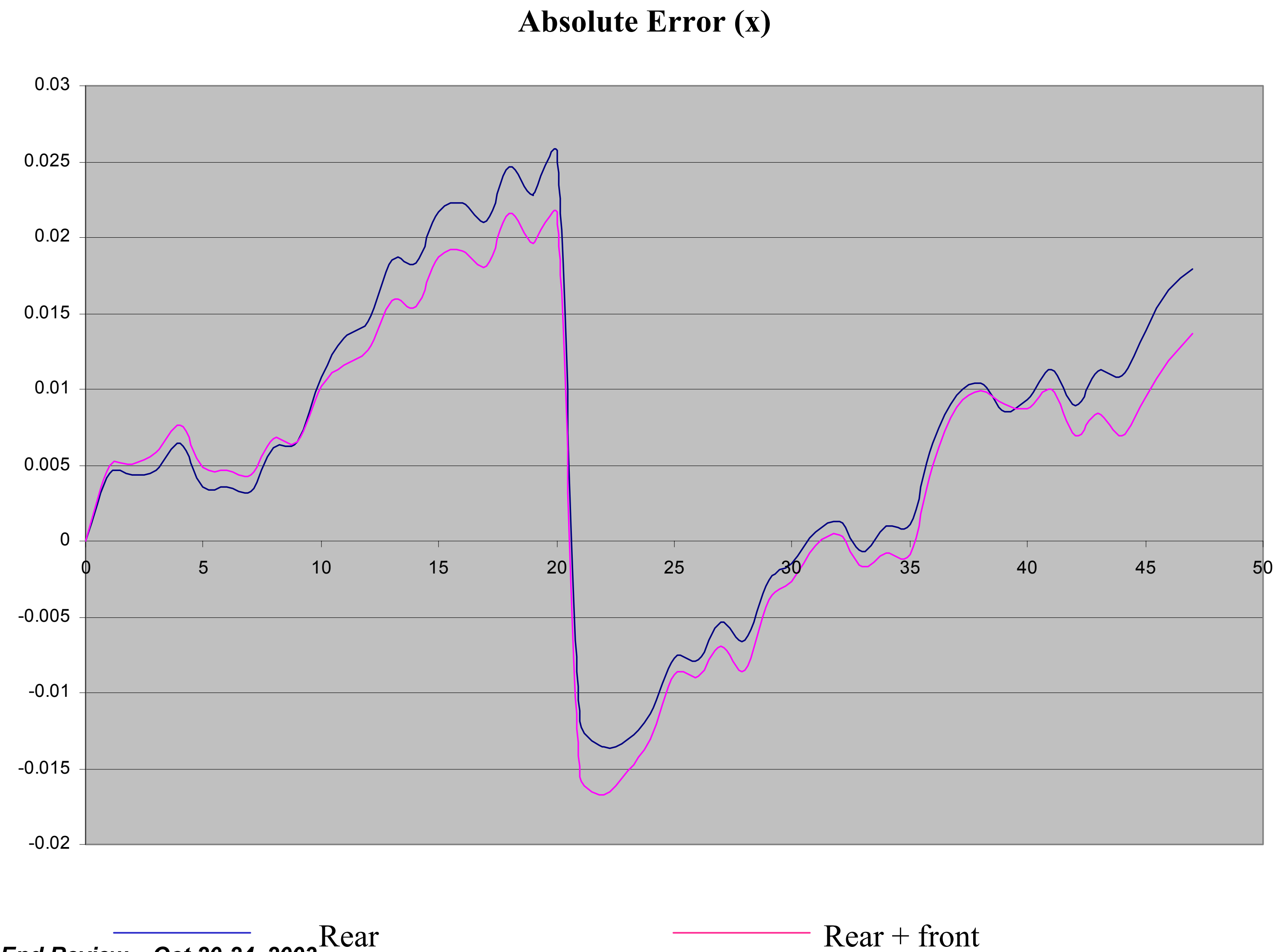
Roll Estimation



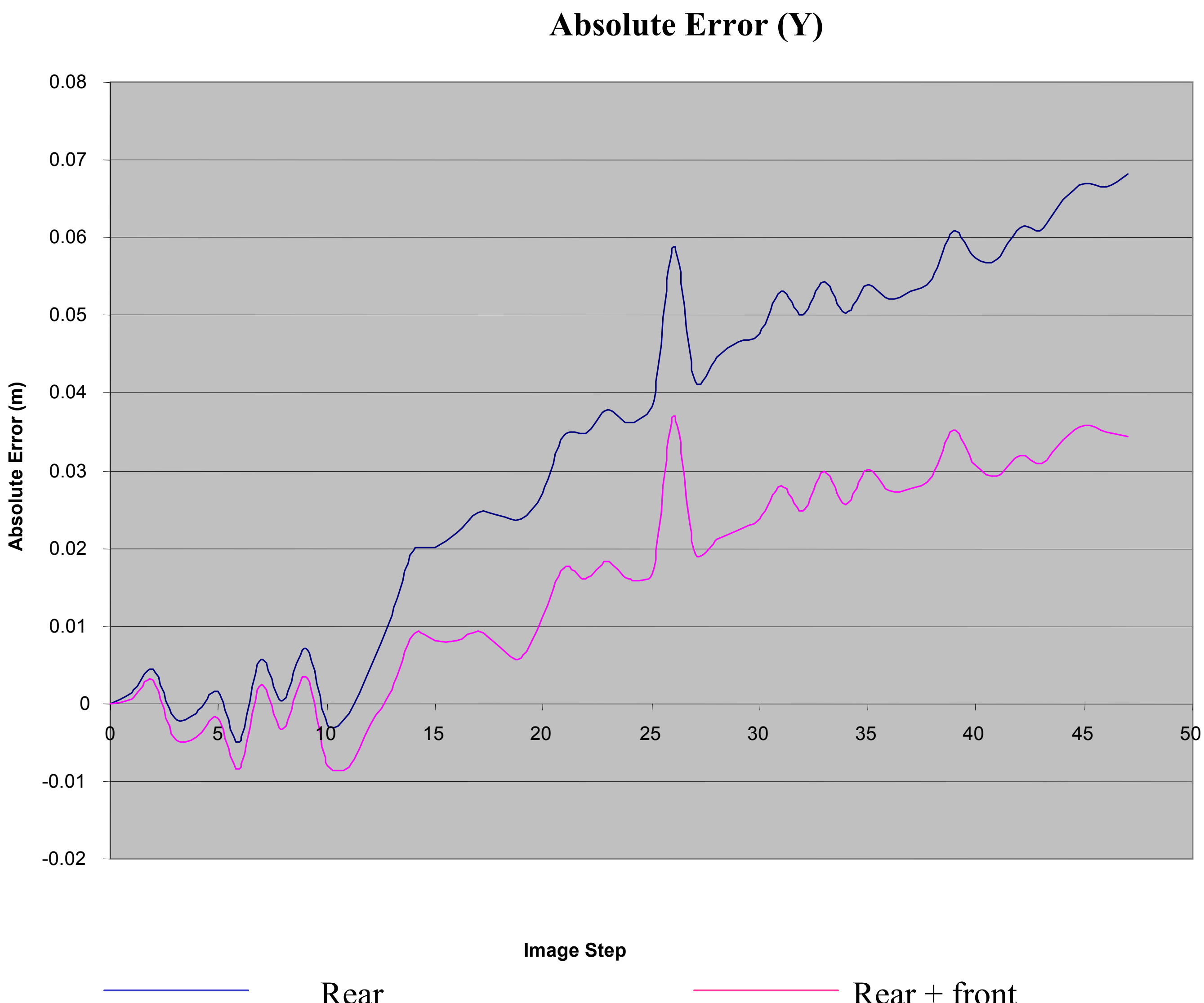
Pitch Estimation



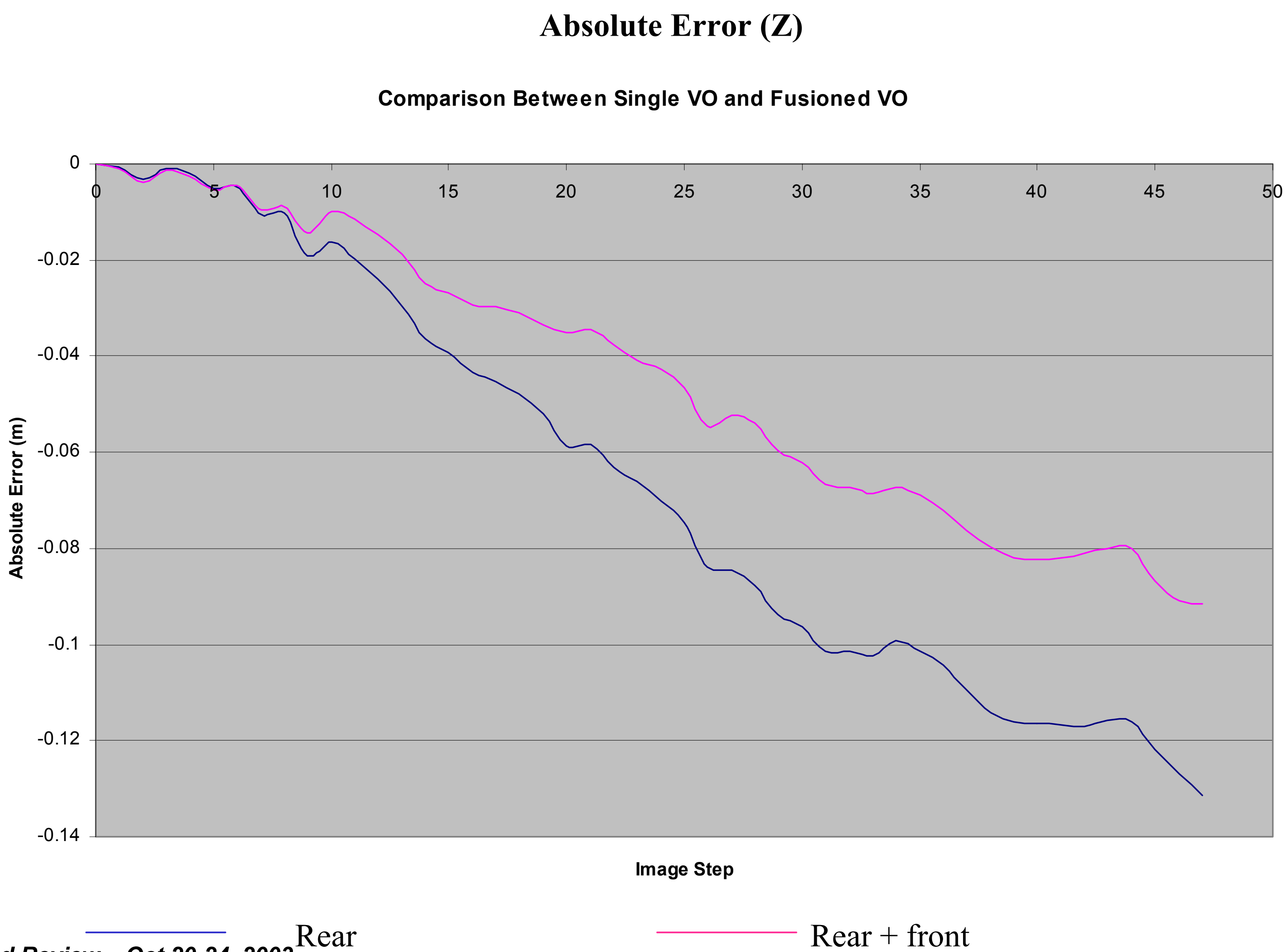
VO Fusion (front and rear Has Camera)



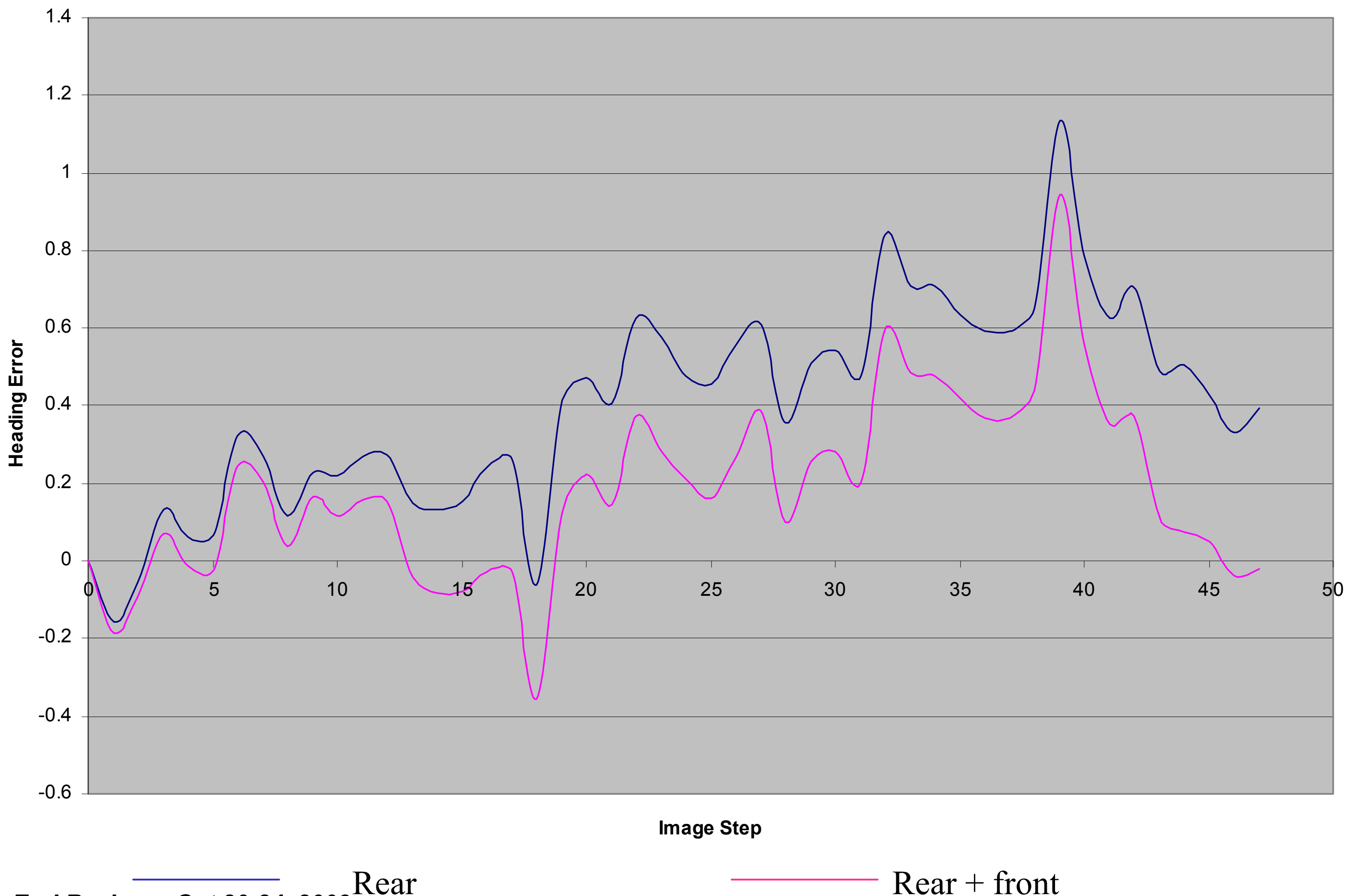
VO Fusion (front and rear Has Camera)



VO Fusion (front and rear Has Camera)



Absolute Error (Heading)



Absolute Error (Pitch)

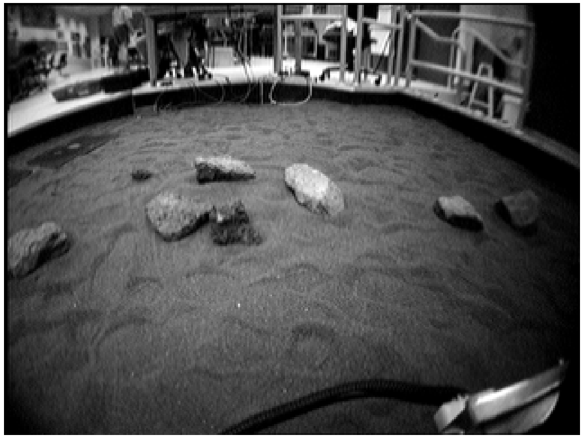
Absolute Error (Roll)

Heading Estimation

0

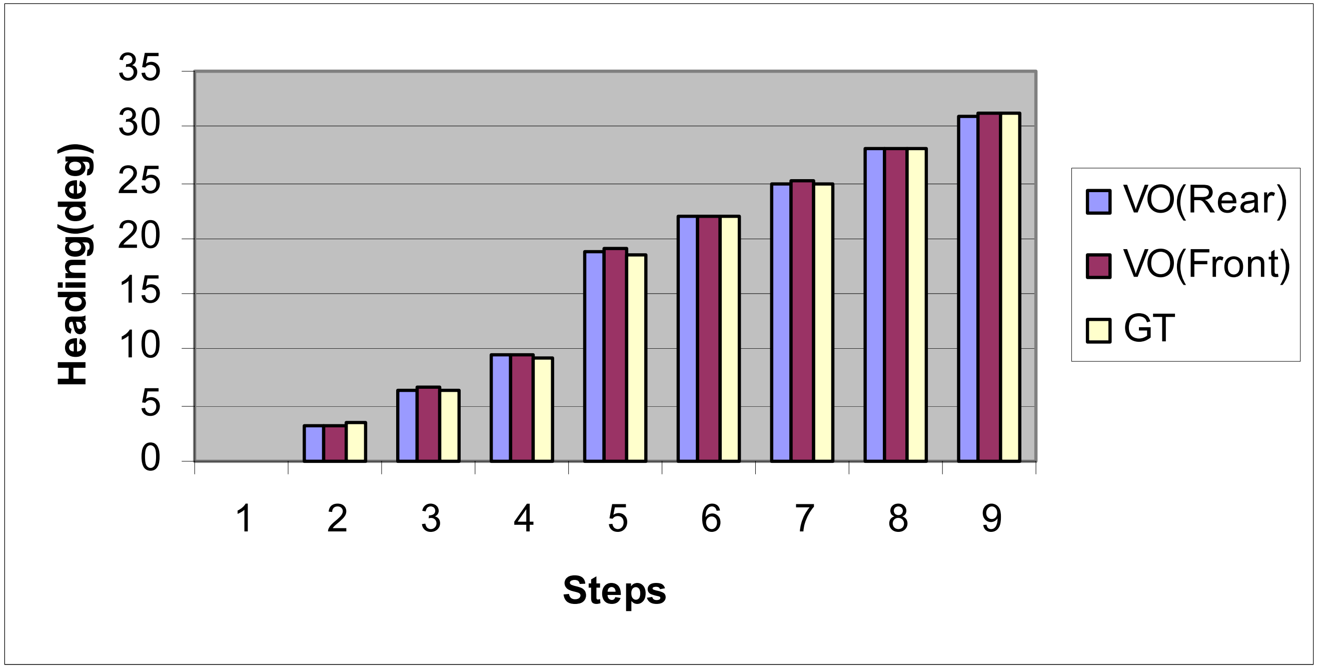
4

7

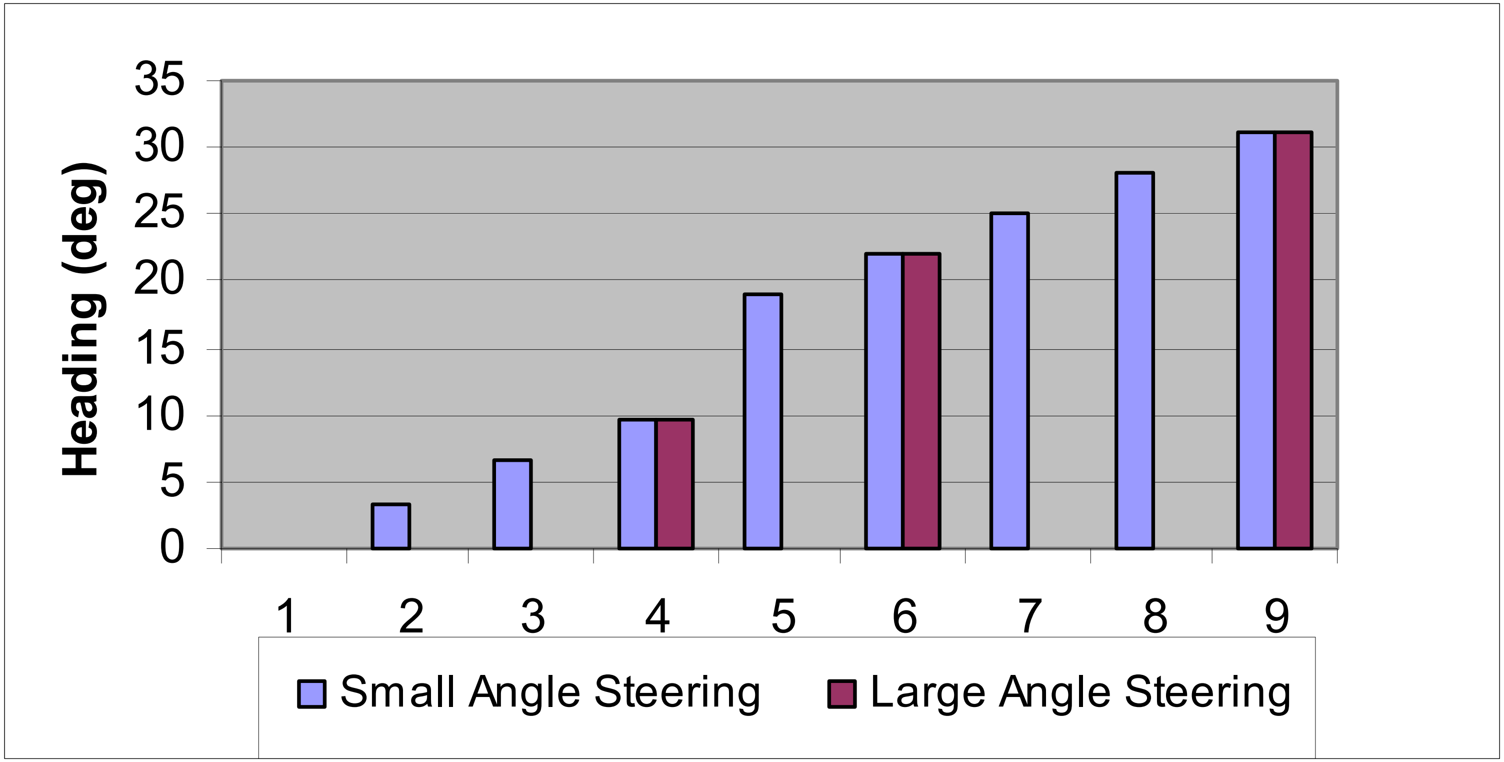


10

Heading Estimation (1)



Heading Estimation (2)



Field Test

<u>Location</u>

- Johnson Valley, Mojave Desert, CA
- Sandy slopes of up to 20-25° slopes

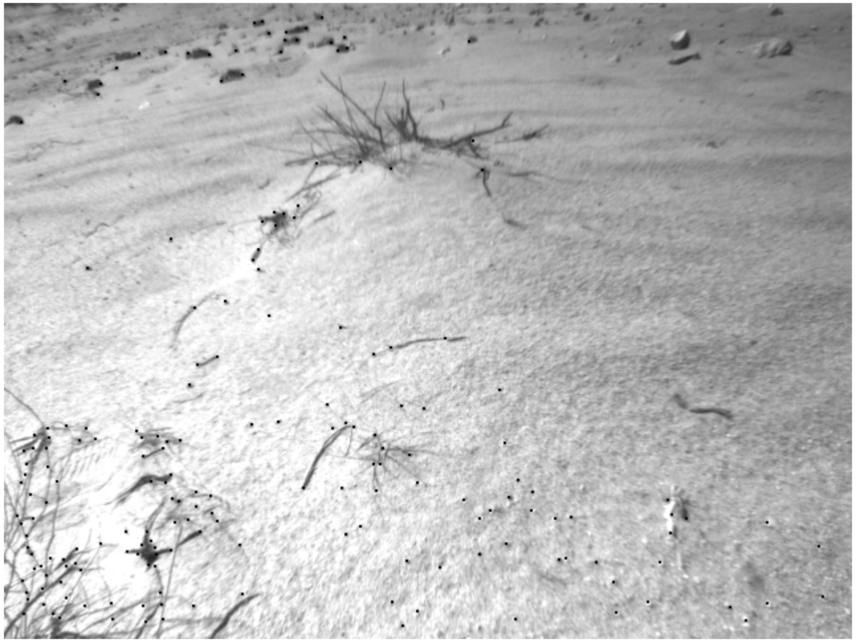
<u>Logistics</u>

- 4 days – 4 people
  - 1.5 days of setup and break down
  - 2.5 days of experimentation

<u>Motivation</u>

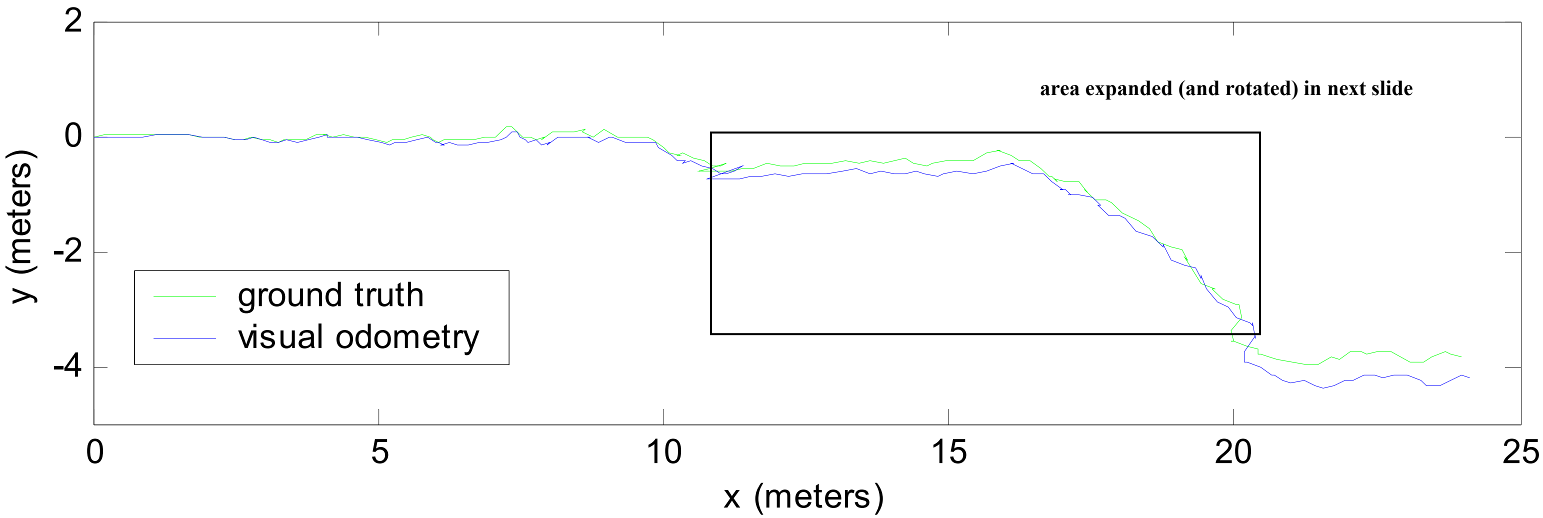
- Mars Yard is too small and has no slopes
  - The size is mostly a factor for visual odometry which looks far beyond traverse distance

Sample of images



Field Test Results

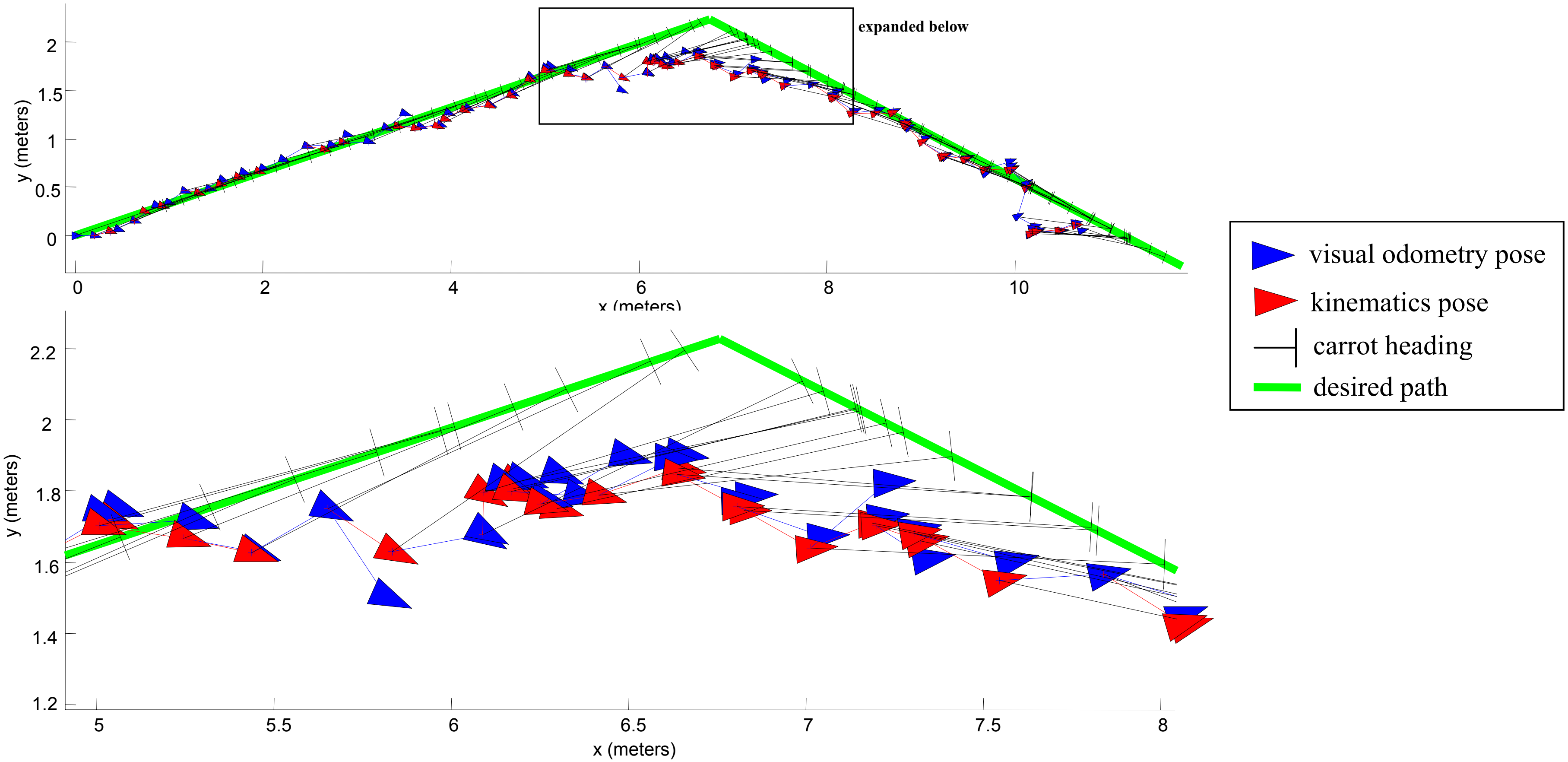
Visual Odometry vs. Ground Truth



- Error (0.37 m) is less than 1.5% of distance traveled (29 m)
- Ground truth data collected with a Leica Total Station and four rover mounted prisms

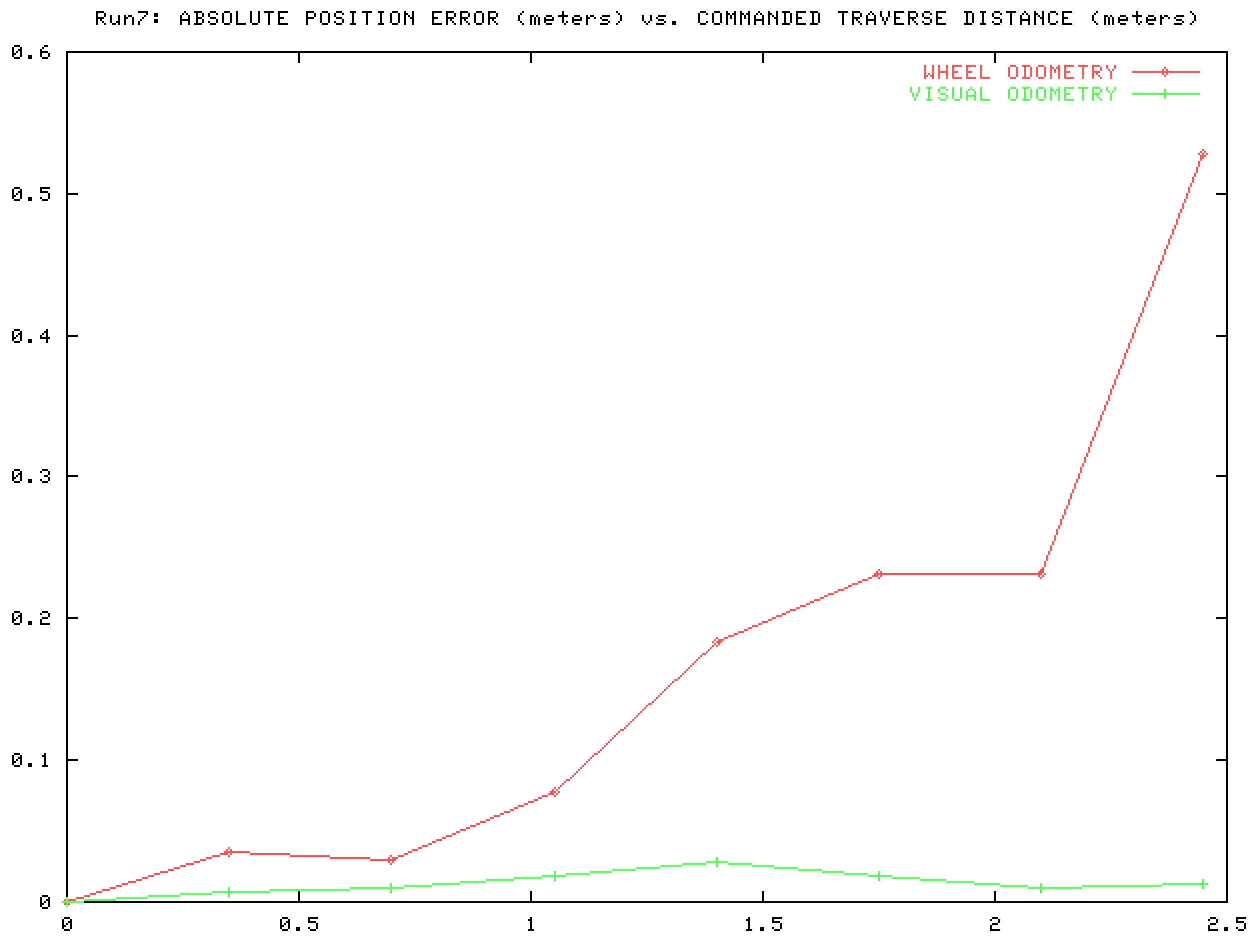
Field Test Results

Slip Compensation/Path Following Results

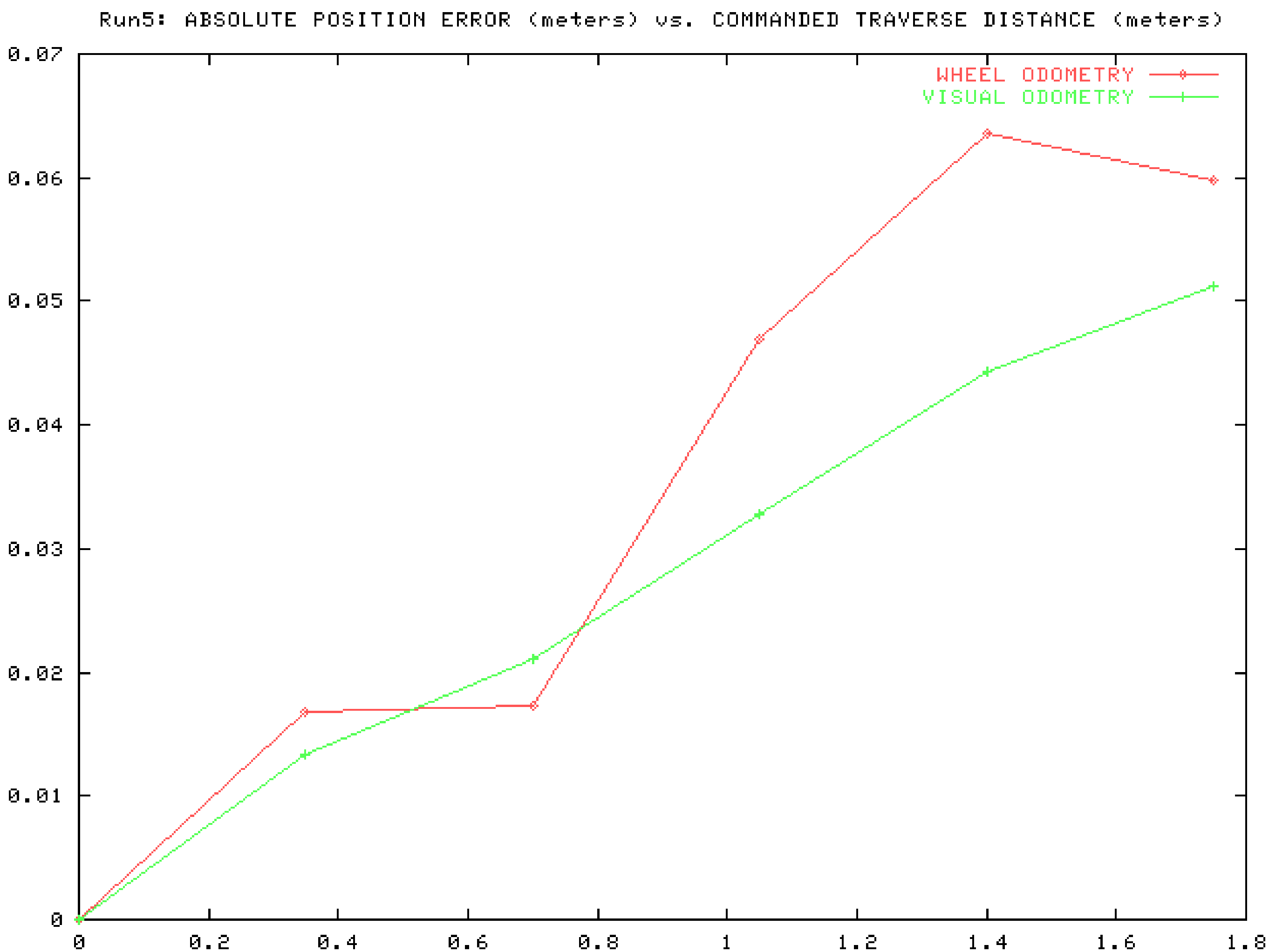


- There is a noticeable bias between the visual odometry pose and the kinematics pose in the y direction of many estimates; this is due to the downhill slippage of the rover; this bias is being compensated for in the slip compensation algorithm

MER VO Test (Rough)

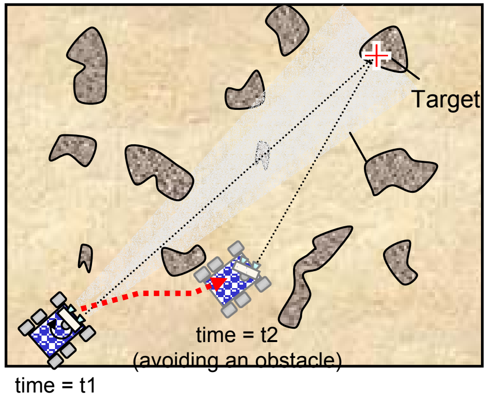


MER Test

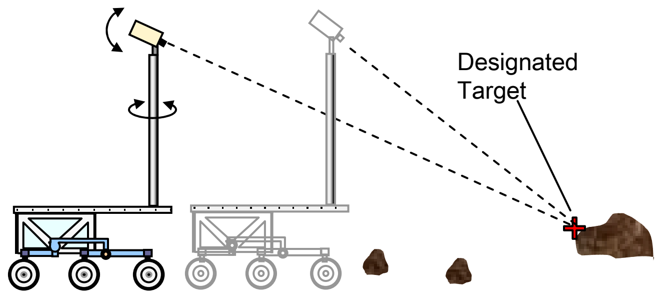


Run5: ABSOLUTE POSITION ERROR (meters) vs. COMMANDED TRAVERSE DISTANCE (meters)
WHEEL ODOMETRY
VISUAL ODOMETRY
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8

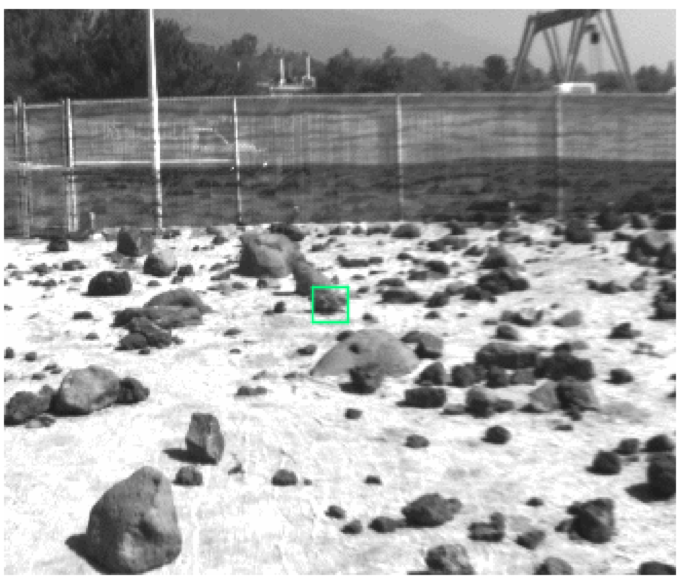
Target Approach



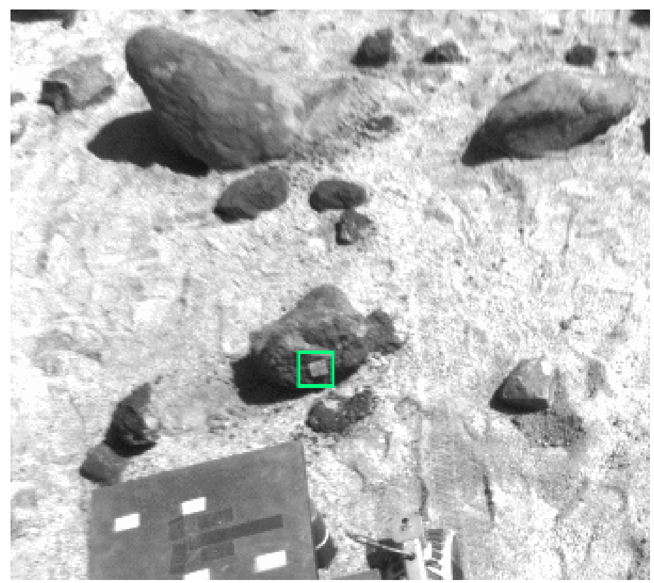
(a)



(b)

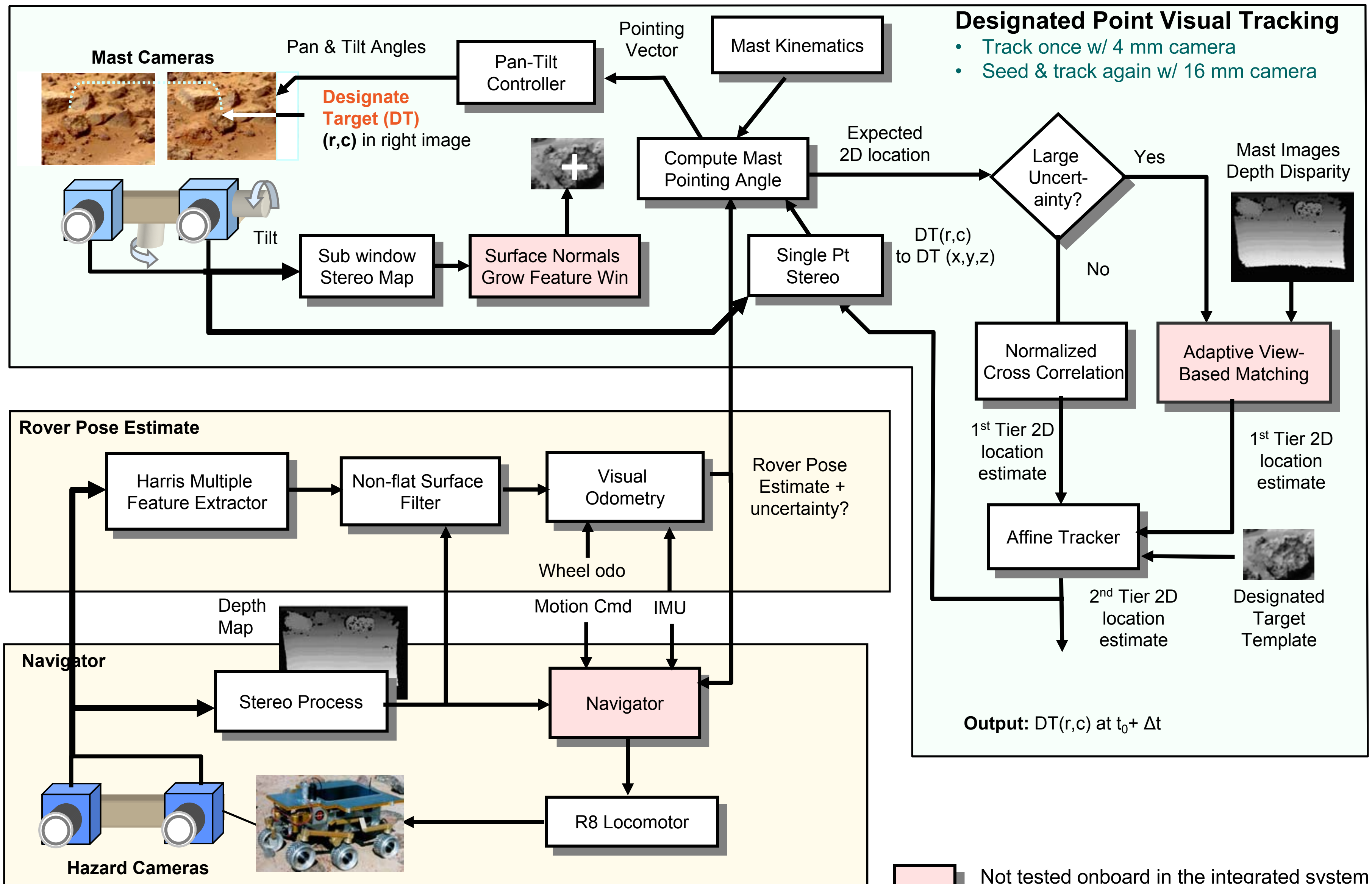


1st Frame



37th Frame after 10 m

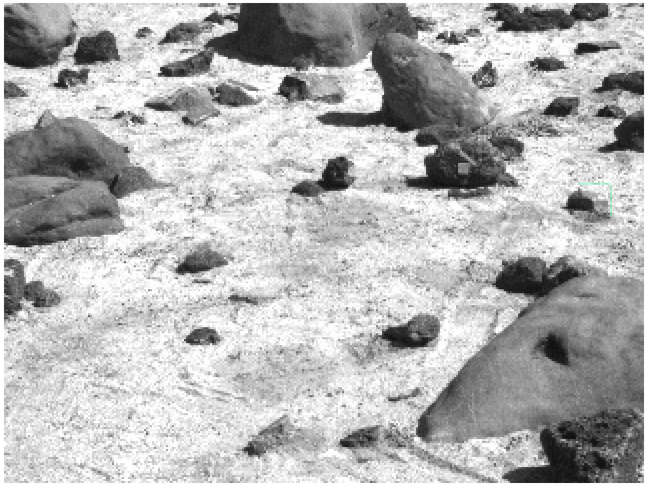
Integrated 2D/3D Tracker



Tracking Results over Rough Terrain

Tracking Video


View from 4 mm camera


View from 16 mm camera

Ground Truth Data Collection System

- Automatically tracks the position of 1 prism and finds the 3 other prisms when rover stops
- Simplifies and speeds the collection of ground truth data in field tests
- Locates rover frame in world frame and the initial rover frame
- +/- 2mm position accuracy
- +/- 0.3º orientation accuracy

Future Works

- A real-time Visual Odometry
- Data Fusion with other sensors (IMU …) to achieve better estimation
- Visual Odometry Applications